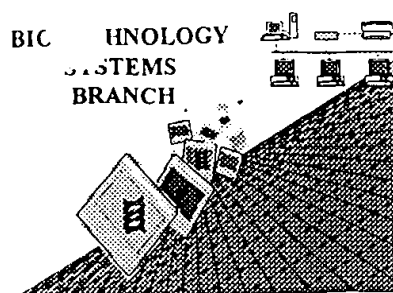


## RAW SEQUENCE LISTING ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/831,142

Source: RET 09

Date Processed by STIC: 5-23-01

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

PCT09

## RAW SEQUENCE LISTING

DATE: 05/23/2001

PATENT APPLICATION: US/09/831,142

TIME: 16:03:42

Input Set : A:\WCM69US.app

Output Set: C:\CRF3\05232001\I831142.raw

3 <110> APPLICANT: University of Wales College of Medicine  
 5 <120> TITLE OF INVENTION: Protein and DNA coding therefor  
 7 <130> FILE REFERENCE: PCT/GB99/03654  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/831,142  
 C--> 10 <141> CURRENT FILING DATE: 2001-05-07  
 12 <160> NUMBER OF SEQ ID NOS: 22  
 14 <170> SOFTWARE: PatentIn Ver. 2.1  
 16 <210> SEQ ID NO: 1  
 17 <211> LENGTH: 870  
 18 <212> TYPE: DNA  
 19 <213> ORGANISM: Pholas dactylus  
 21 <400> SEQUENCE: 1  
 22 gaattcggca cgagtcggaa aagaacaaaa tggcttgtat cgttttcggt gctcttgtcg 60  
 23 ctctatgctt aatgcaaccg ggttcgggtg aggaagtaca atgcgcgatg aattggacac 120  
 24 aagctaataa atatgtgttc aacgtggact ggatgaccat ttcatctac gactatggcg 180  
 25 ctcaagagca actgtacgaa gatcgggctt tggggctgtg tcggattgaa cgggccggcc 240  
 26 caggtaccac aaaagccgtc tggattaact ggagtaacga cacgcagtca tgtgtaacaa 300  
 27 gaaaaacaat cttcttcgag gttggtggag aaattgcccg gctagttgac tacagaccac 360  
 28 aggaagacgg aactgagaaa actttttacaa gaaaattctc tagcaaatg ccaggcactt 420  
 29 acatgcttat ggacgtgtgc gctacaaggg acgctgatga taaatgcac gaaggcaca 480  
 30 ttgtggtgac agtcagggtg tccctatatg acgaagataa caatggtgta atggatgaag 540  
 31 gtaagggtgat tccatctgag acaatcgagg atgatataa ggactgtggg ctcttagacc 600  
 32 aagatgttga actcgattat acgtggactc aaaacgagtg tgatctacca gacacagtag 660  
 33 acgaggctga agacacaccg tcagaaactg gagaattctt ctggtagatc tatcagacta 720  
 34 cttttatcag caggacaact ggtcgttacc agacacctat aacgtgtcct catcaataat 780  
 35 gtgtaaaaca gaaataatcg atagaatatt gaaaataaaa tgtaataaaa cactgggtga 840  
 36 aatatgaaaa aaaaaaaaaa aaaactcgag 870  
 39 <210> SEQ ID NO: 2  
 40 <211> LENGTH: 816  
 41 <212> TYPE: DNA  
 42 <213> ORGANISM: Pholas dactylus  
 44 <400> SEQUENCE: 2  
 45 gaattcggca cgagggaaaa gaacaaaatg gcttgtatcg ttttcggtgc tcttgtcgct 60  
 46 ctatgcttaa tgcaaccggg ttccgggtgag gaagtacaat gcgcgatgaa ttggacacaa 120  
 47 gctaataaat atgtgttcaa cgtggactgg atgaccattt tcatctacga ctatggcgct 180  
 48 caagagcaac tgtaacagga tcgggctttg gggctgtgtc ggattgaacg ggccggccca 240  
 49 ggtaccacaa aagccgtctg gattaactgg agtaacgaca cgcagtcacg tgtaacaaga 300  
 50 aaaacaatct tcttcgaggt tgggtggagaa attgcccggc tagttgacta cagaccacag 360  
 51 gaagacggaa ctgagaaaaa tttttacaaga aaattctcta gcaaatgcc aggcacttac 420  
 52 atgcttatgg acgtgtgcgc tacaagggac gctgatgata aatgcacga aggcacaatt 480  
 53 gtggtgacag tcagggtgtc cctatatgac gaagataaca atggtgtaat ggatgaagg 540  
 54 aaggttattc catctgagac aatcgaggat gatatacaagg actgtgggct cttagaccaa 600  
 55 gatgttgaac tcgattatac gtggactcaa aacgagtgtg atctaccaga cacagtagac 660  
 56 gaggtgaag acacaccgtc agaaactgga gaattcttct ggtagatcta tcagaccact 720  
 57 tttatcagca ggacaactgg tcgttaccag acacctataa cgtgtcctca tcaataatgt 780  
 58 gtaaacacaga aataatcgat agaattatga aaataa 816  
 61 <210> SEQ ID NO: 3

Does Not Comply  
 Corrected Diskette Needed  
 pp 4, 5

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,142

DATE: 05/23/2001

TIME: 16:03:42

Input Set : A:\WCM69US.app

Output Set: C:\CRF3\05232001\I831142.raw

```

62 <211> LENGTH: 852
63 <212> TYPE: DNA
64 <213> ORGANISM: Pholas dactylus
66 <400> SEQUENCE: 3
67 gtcggaag aacaaatgg cttgtatcgt tttcgttgct cttgtcgtc tatgcttaat 60
68 gcaaccgggt tccggtgagg aagtacaatg cgcgatgaat tggacacaag ctaatgaata 120
69 tgtgttcaac gtggactgga tgaccatttt catctacgac tatggcgctc aagagcaact 180
70 gtacgaggat cgggctttgg ggctgtgtcg gattgaacgg gccggcccag gtaccacaaa 240
71 agccgtctgg attaaactgga gtaacgacac gcagtcatgt gtaacaagaa aaacaatctt 300
72 cttcgagggt ggtggagaaa ttgcccggt agttgactac agaccacagg aagacggaac 360
73 tgagaaaact ttacaagaa aattctctag caaaatgccca ggcaattaca tgcttatgga 420
74 cgtgtgcgct acaagggacg ctgatgataa atgcatcgaa ggcacaattg tggtgacagt 480
75 cagggtgtcc ctatatgacg aagataacaa tgggtgtaat gatgaaggta aggttattcc 540
76 atctgagaca atcgaggatg atatcaagga ctgtgggctc ttagaccaag atgttgaact 600
77 cgattatacg tggactcaaa acgagtgtga tctaccagac acagtagacg aggttgaaga 660
78 cacaccgtca gaaactggag aattcttctg gtagatctat cagaccactt ttatcagcag 720
79 gacaactggg cgttaccaga cacctataac gtgtcctcat caataatgtg taaaacagaa 780
80 ataatcgata gaatattgaa aataaaatgt taatagacac tggttgaaaa aaaaaaaaaa 840
81 aaaaaactcg ag 852
84 <210> SEQ ID NO: 4
85 <211> LENGTH: 225
86 <212> TYPE: PRT
87 <213> ORGANISM: Pholas dactylus
89 <400> SEQUENCE: 4
90 Met Ala Cys Ile Val Phe Val Ala Leu Val Ala Leu Cys Leu Met Gln
91 1 5 10 15
93 Pro Gly Ser Gly Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala
94 20 25 30
96 Asn Glu Tyr Val Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp
97 35 40 45
99 Tyr Gly Ala Gln Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys
100 50 55 60
102 Arg Ile Glu Arg Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn
103 65 70 75 80
105 Trp Ser Asn Asp Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe
106 85 90 95
108 Glu Val Gly Gly Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu
109 100 105 110
111 Asp Gly Thr Glu Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro
112 115 120 125
114 Gly Thr Tyr Met Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp
115 130 135 140
117 Lys Cys Ile Glu Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr
118 145 150 155 160
120 Asp Glu Asp Asn Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser
121 165 170 175
123 Glu Thr Ile Glu Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp
124 180 185 190
126 Val Glu Leu Asp Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp

```

## RAW SEQUENCE LISTING

DATE: 05/23/2001

PATENT APPLICATION: US/09/831,142

TIME: 16:03:42

Input Set : A:\WCM69US.app

Output Set: C:\CRF3\05232001\I831142.raw

```

127          195          200          205
129 Thr Val Asp Glu Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe
130      210          215          220
132 Trp
133 225
136 <210> SEQ ID NO: 5
137 <211> LENGTH: 205
138 <212> TYPE: PRT
139 <213> ORGANISM: Pholas dactylus
141 <400> SEQUENCE: 5
142 Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala Asn Glu Tyr Val
143   1          5          10          15
145 Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp Tyr Gly Ala Gln
146      20          25          30
148 Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys Arg Ile Glu Arg
149      35          40          45
151 Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn Trp Ser Asn Asp
152      50          55          60
154 Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe Glu Val Gly Gly
155  65          70          75          80
157 Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu Asp Gly Thr Glu
158      85          90          95
160 Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro Gly Thr Tyr Met
161      100          105          110
163 Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp Lys Cys Ile Glu
164      115          120          125
166 Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr Asp Glu Asp Asn
167      130          135          140
169 Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser Glu Thr Ile Glu
170 145          150          155          160
172 Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp Val Glu Leu Asp
173      165          170          175
175 Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp Thr Val Asp Glu
176      180          185          190
178 Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe Trp
179      195          200          205
182 <210> SEQ ID NO: 6
183 <211> LENGTH: 225
184 <212> TYPE: PRT
185 <213> ORGANISM: Pholas dactylus
187 <400> SEQUENCE: 6
188 Met Ala Cys Ile Val Phe Val Ala Leu Val Ala Leu Cys Leu Met Gln
189   1          5          10          15
191 Pro Gly Ser Gly Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala
192      20          25          30
194 Asn Glu Tyr Val Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp
195      35          40          45
197 Tyr Gly Ala Gln Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys
198      50          55          60

```

## RAW SEQUENCE LISTING

DATE: 05/23/2001

PATENT APPLICATION: US/09/831,142

TIME: 16:03:42

Input Set : A:\WCM69US.app

Output Set: C:\CRF3\05232001\I831142.raw

```

200 Arg Ile Glu Arg Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn
201 65 70 75 80
203 Trp Ser Asn Asp Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe
204 85 90 95
206 Glu Val Gly Gly Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu
207 100 105 110
209 Asp Gly Thr Glu Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro
210 115 120 125
212 Gly Thr Tyr Met Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp
213 130 135 140
215 Lys Cys Ile Glu Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr
216 145 150 155 160
218 Asp Glu Asp Asn Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser
219 165 170 175
221 Glu Thr Ile Glu Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp
222 180 185 190
224 Val Glu Leu Asp Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp
225 195 200 205
227 Thr Val Asp Glu Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe
228 210 215 220

```

230 Trp

231 225

234 &lt;210&gt; SEQ ID NO: 7

235 &lt;211&gt; LENGTH: 17

236 &lt;212&gt; TYPE: DNA

237 &lt;213&gt; ORGANISM: Pholas dactylus

239 &lt;220&gt; FEATURE:

240 &lt;221&gt; NAME/KEY: modified\_base

241 &lt;222&gt; LOCATION: (3)

242 &lt;223&gt; OTHER INFORMATION: i

244 &lt;400&gt; SEQUENCE: 7

W--&gt; 245 acnathhtyt tycargt

17

248 &lt;210&gt; SEQ ID NO: 8

249 &lt;211&gt; LENGTH: 17

250 &lt;212&gt; TYPE: DNA

251 &lt;213&gt; ORGANISM: Pholas dactylus

253 &lt;220&gt; FEATURE:

254 &lt;221&gt; NAME/KEY: modified\_base

255 &lt;222&gt; LOCATION: (15) /

256 &lt;223&gt; OTHER INFORMATION: i

258 &lt;400&gt; SEQUENCE: 8

W--> 259 cargargarg ~~macn~~ga

17

262 &lt;210&gt; SEQ ID NO: 9

263 &lt;211&gt; LENGTH: 17

264 &lt;212&gt; TYPE: DNA

265 &lt;213&gt; ORGANISM: Pholas dactylus

267 &lt;220&gt; FEATURE:

268 &lt;221&gt; NAME/KEY: modified\_base

269 &lt;222&gt; LOCATION: (3) /

## RAW SEQUENCE LISTING

DATE: 05/23/2001

PATENT APPLICATION: US/09/831,142

TIME: 16:03:42

Input Set : A:\WCM69US.app

Output Set: C:\CRF3\05232001\I831142.raw

270 <223> OTHER INFORMATION: i  
 272 <400> SEQUENCE: 9  
*"A" at position 6?*

W--> 273 ~~tcngtnc~~cyt cytcytg 17  
 276 <210> SEQ ID NO: 10  
 277 <211> LENGTH: 18  
 278 <212> TYPE: DNA  
 279 <213> ORGANISM: Pholas dactylus  
 281 <220> FEATURE:  
 282 <221> NAME/KEY: modified\_base  
 283 <222> LOCATION: (9) /  
 284 <223> OTHER INFORMATION: i  
 286 <400> SEQUENCE: 10

W--> 287 tt~~yaay~~gtng aytggatg 18  
 290 <210> SEQ ID NO: 11  
 291 <211> LENGTH: 20  
 292 <212> TYPE: DNA  
 293 <213> ORGANISM: Pholas dactylus  
 295 <400> SEQUENCE: 11  
 296 acacagcccc aaagcccgat 20  
 299 <210> SEQ ID NO: 12  
 300 <211> LENGTH: 20  
 301 <212> TYPE: DNA  
 302 <213> ORGANISM: Pholas dactylus  
 304 <400> SEQUENCE: 12  
 305 ttgcccggct agttgactac 20  
 308 <210> SEQ ID NO: 13  
 309 <211> LENGTH: 24  
 310 <212> TYPE: DNA  
 311 <213> ORGANISM: Pholas dactylus  
 313 <400> SEQUENCE: 13  
 314 catatttcaa ccagtgttta ttaa 24  
 317 <210> SEQ ID NO: 14  
 318 <211> LENGTH: 19  
 319 <212> TYPE: DNA  
 320 <213> ORGANISM: Pholas dactylus  
 322 <400> SEQUENCE: 14  
 323 caattgtgcc ttcgatgca 19  
 326 <210> SEQ ID NO: 15  
 327 <211> LENGTH: 17  
 328 <212> TYPE: DNA  
 329 <213> ORGANISM: Pholas dactylus  
 331 <400> SEQUENCE: 15  
 332 ggactgtggg ctcttag 17  
 335 <210> SEQ ID NO: 16  
 336 <211> LENGTH: 20  
 337 <212> TYPE: DNA  
 338 <213> ORGANISM: Pholas dactylus  
 340 <400> SEQUENCE: 16  
 341 atggcttgta tcgttttcgt 20

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/831,142

DATE: 05/23/2001

TIME: 16:03:43

Input Set : A:\WCM69US.app

Output Set: C:\CRF3\05232001\I831142.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:245 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7

L:259 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8

L:273 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9

L:287 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10